Executive summary

The IFIPDT examination of the two WRF cores generally showed little statistical difference in the performance of the cores. However, close examination of case studies reveals that smaller-scale features are more readily evident in the ARW, while they appear to be smoother in the NMM. This is also borne out in distributions of the frequency of occurrence of certain parameters, both when icing is present and overall. IFIPDT experience is that icing conditions occur on a wide range of scales, from hundreds of meters to hundreds of kilometers, with much of the variability on the smaller end of that range. Though we cannot validate that the smaller scales and stronger vertical velocities depicted by the ARW are real, we recognize that these scales are important for icing and have hope that we may be able to make use of them in IFIPDT systems, should they prove to be valid. Thus, the IFIPDT has a slight preference for the ARW core. Regardless of which core is chosen, the IFIPDT will need to tune, calibrate, revise our use of certain fields, and/or determine ways to handle the scales that dominate the model output used in FIP and CIP.

Introduction

The In-Flight Icing Product Development Team (IFIPDT) created the Current and Forecast Icing Products (CIP and FIP) to diagnose and forecast icing conditions and their severity across the RUC and Alaska NAM domains. These products make use of model forecasts of temperature (T), relative humidity (RH), explicit condensate fields (cloud, rain, ice, snow and graupel water content), vertical velocity (VV), and precipitation. FIP depends solely on these fields to produce forecasts of icing out to 2, 3, 6, 9 and 12 hours, using the RUC hybrid-b grids. CIP merges model output with observations from satellite, radar, surface stations, the NLDN and pilot reports to produce hourly diagnoses of icing. This study focused on the T, RH, condensate and VV fields and their potential downstream effects on FIP. FIP was chosen since model fields provide the sole input to the product suite and the effects of differences in the model output would not be masked by the observational data employed by CIP.

WRF model output from both cores (ARW and NMM) for both phases (1 and 2) were examined for a set of sixty 12-h forecasts and sixty 6-h forecasts, with all runs initiated at 1200 UTC. These runs were chosen to have valid times where a significant number of pilot reports (PIREPs) of icing would be available for comparison and in an attempt to eliminate any moisture issues that could be associated with the first few hours of the forecast cycle. Core-to-core comparisons were the primary focus of the study, since phase-to-phase comparisons are slightly confounded by differences in other aspects of the runs (e.g. convective and boundary layer schemes). However, some differences between the two microphysics packages (Ferrier in phase 1 and Thompson in phase 2) and between WRF and RUC output will be noted where important.

Statistical verification

In this section, we compare PIREPs of icing to the primary WRF output fields that would contribute to FIP icing forecasts. This was done for the 120 output times described above. While it would have be ideal to have run FIP on these 480 WRF files (120 valid times, 2 cores, 2 phases), it was not practical for this experiment. Instead, output from the explicit microphysics packages, were verified against PIREPs. In addition, distributions of T, RH, condensate and VV were made for the PIREPs, as well as for the model grids, regardless of the presence of PIREPs (e.g. RH across the domain at all times and altitudes). The team was able to create some WRF-based FIP runs for the purpose of visual inspection, but there were not enough runs to create meaningful statistics.

Using methods similar to those described by Brown et al (1997), WRF explicit forecasts of supercooled liquid water (cloud and rain water; hereafter referred to as "SLW") and the sum of all condensate fields (cloud, rain, ice and snow water; total condensate, hereafter referred to as "TotC") at $T<0^{o}C$ were matched to both positive and explicit negative PIREPs of icing. For PIREPs that occurred within one hour of the WRF valid time, their latitude and longitude were mapped to the model grid, then the nine points surrounding the PIREP were examined. The PIREP altitude was used to find the closest vertical levels above and below. From this set of 18 points, the maximum value of SLW and TotC was calculated. For 0000 UTC valid times, a total of 1097 positive and 338 negative PIREPs (1305 positive and 578 negative PIREPs for 1800 UTC valid times) were compared to the WRF 12-h (6-h) forecasts. Very low thresholds of SLW and TotC were used to diagnose a "yes" forecast. If any condensate in these categories was predicted, the forecast was a "yes". Results are shown in Tables 1 and 2.

Table 1: Probability of detection (POD) of positive (y) and negative (n) icing PIREPs from 12-h and 6-h forecasts valid at 0000 and 1800 UTC.

SLW	**ARW-1**	**NMM-1**	**ARW-2**	**NMM-2**	**RUC**
PODy (12h)	0.24	0.22	0.48	0.47	0.39
PODn (12h)	0.90	0.92	0.82	0.82	0.86
PODy (6h)	0.22	0.23	0.43	0.43	0.39
PODn (6h)	0.85	0.84	0.81	0.80	0.87
TotC	**ARW-1**	**NMM-1**	**ARW-2**	**NMM-2**	**RUC**
PODy (12h)	0.74	0.73	0.69	0.72	0.68
PODn (12h)	0.74	0.73	0.70	0.69	0.74
PODy (6h)	0.72	0.73	0.70	0.68	0.66
PODn (6h)	0.66	0.67	0.72	0.72	0.80

Overall, there was very little core-to-core difference in the verification, regardless of the phase of the test. Both cores predicted the presence of condensate for about 70% of all positive PIREPs. Little core-to-core difference was found for the phase of condensate predicted at these locations. However, phase 1 tests indicate that model produced ice-phase condensate for about 2/3 of these PIREPs and SLW for only 1/3 of them. These results were roughly reversed for phase 2 tests, where roughly 2/3 of the PIREPs with condensate had SLW and only 1/3 had ice.

Distributions

Although the statistical verification showed very little core-to-core difference in the condensate fields, an examination of the distributions of FIP-relevant fields to PIREPs and for the model grids as a whole did reveal some differences. For the sake of brevity, only 12-h forecast output will be discussed here, since results for the 6-h forecasts were very similar. Each chart in this section shows the normalized distribution of counts for each parameter, placed into bins. The maximum normalized value of 1.0 represents the maximum number of counts in any bin, with all other values normalized to that count. For example, if 100 PIREPs had temperatures that fell into the -8.999 to -8.0 temperature bin, and that was the largest count for all bins, the normalized value would be 1.0 for that bin. If the -9.999 to -9.0 bin had 80 PIREPs matched to it, the normalized value would be 0.8. Note that the distributions for the RH and condensate fields were limited to the range of temperatures where FIP can predict icing (–25 to 0^{o}C).

The temperature distributions for all positive icing PIREPs and for the entire model grids (Figs. 1a,b) show that the distributions are essentially the same in a core-to-core sense. No discernable difference was present when looking at positive PIREPs, but the NMM may be a bit cooler in the overall model distributions. The difference is small. We cannot explain the large discrepancy between the overall T distribution for the RUC when compared to the WRF, but suspect that this is attributable to differences in the vertical grid structure perhaps capturing more levels in the colder, drier upper atmosphere in the WRF.

Relative humidity distributions for positive icing PIREPs are quite similar between the two cores (Fig. 2a), though the overall distributions differ somewhat (Fig. 2b). In general, the ARW appears to be wetter than the NMM, with a higher relative frequency of RH in the higher bins. This trend is present in both phases. The two phases also have significant differences, with a pronounced peak in the overall RH around 80-90% in the phase 1 (Ferrier) runs, while only a weaker secondary peak is present in the 85-90% bin in the phase 2 (Thompson) runs. The phase 2 runs have a much stronger peak in the 95-100% bin. This apparently is due to a difference in the triggering point (RH thresholds) of the microphysical species between the two packages. When looking at positive PIREPs, there is little discernable core-to-core differences, though the ARW has a few more occurrences of RH=95-100% than the NMM. That difference is likely not to be statistically significant, given the number of PIREPs available for this analysis. The more striking difference is between the two phases, with phase 2 runs being more skewed toward the highest RH bins.

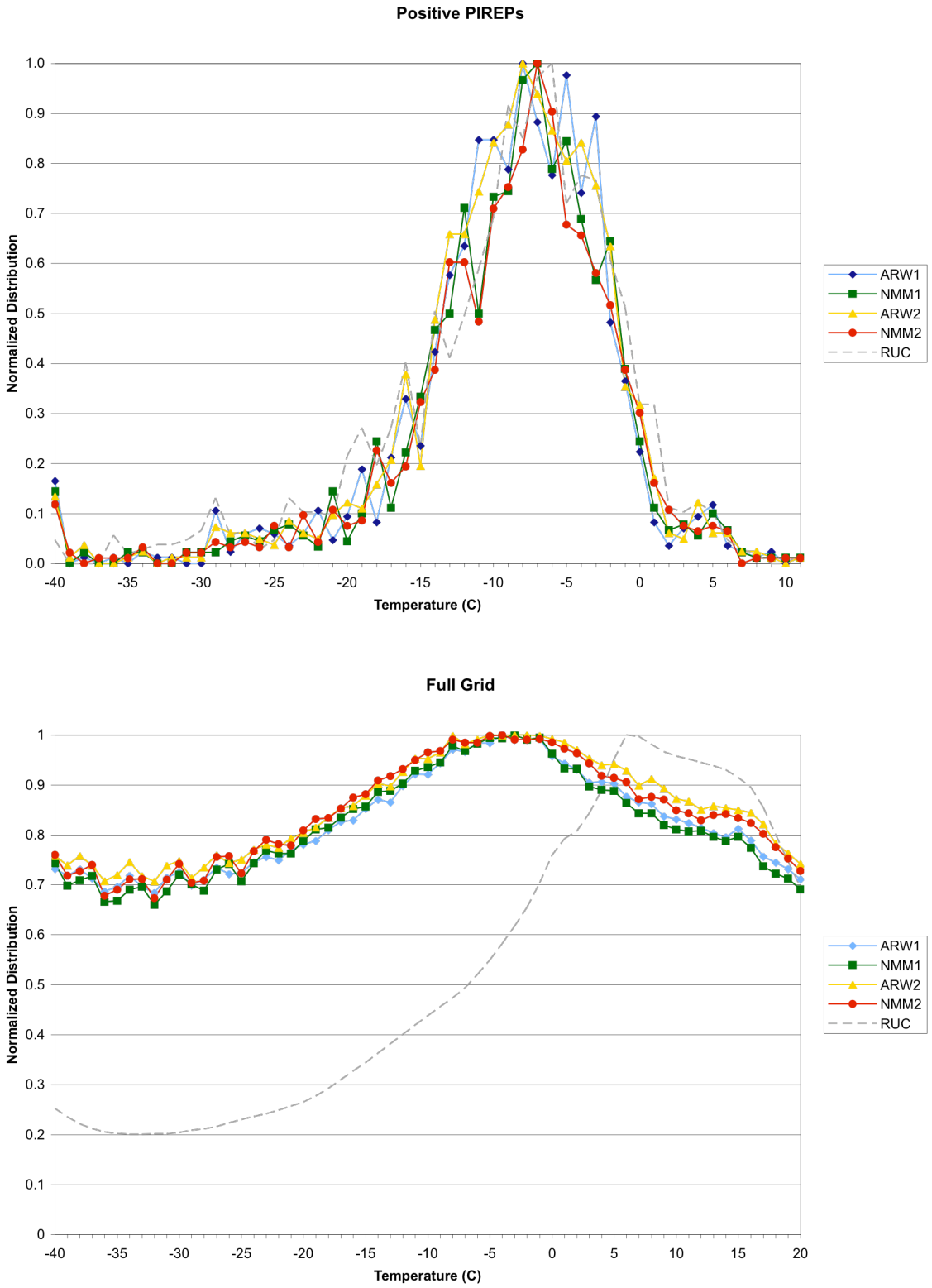
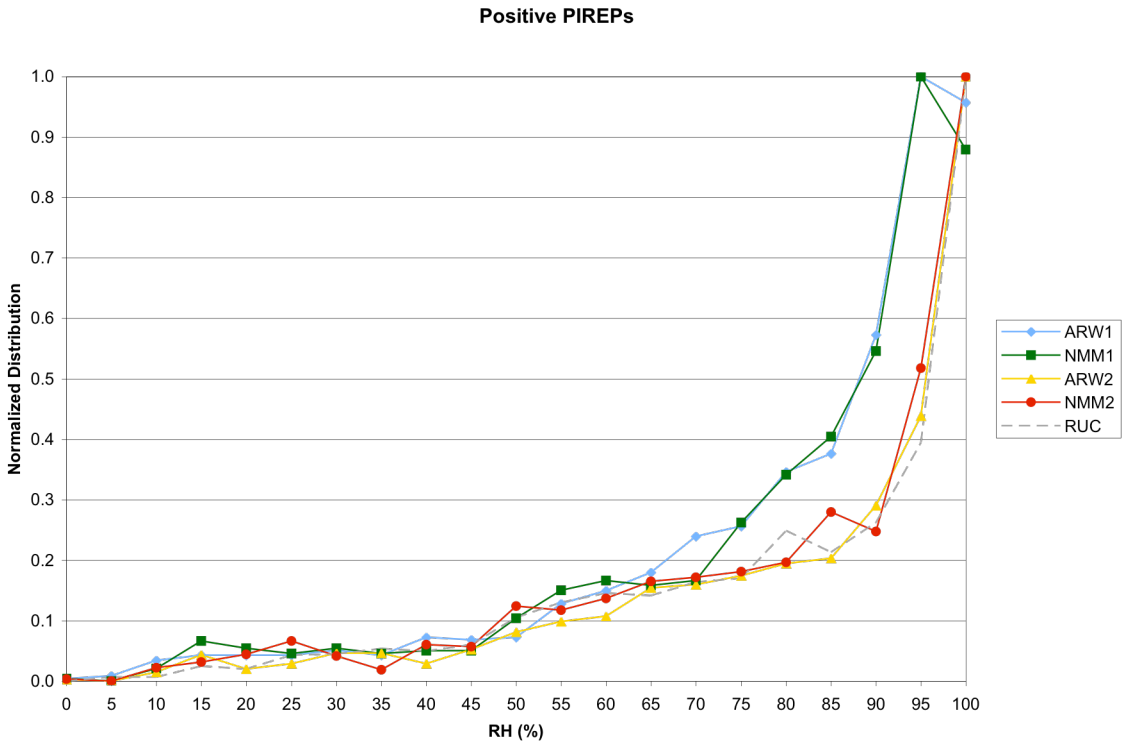


Fig. 1 – T distributions for a) positive icing PIREPs and b) the full model grid.



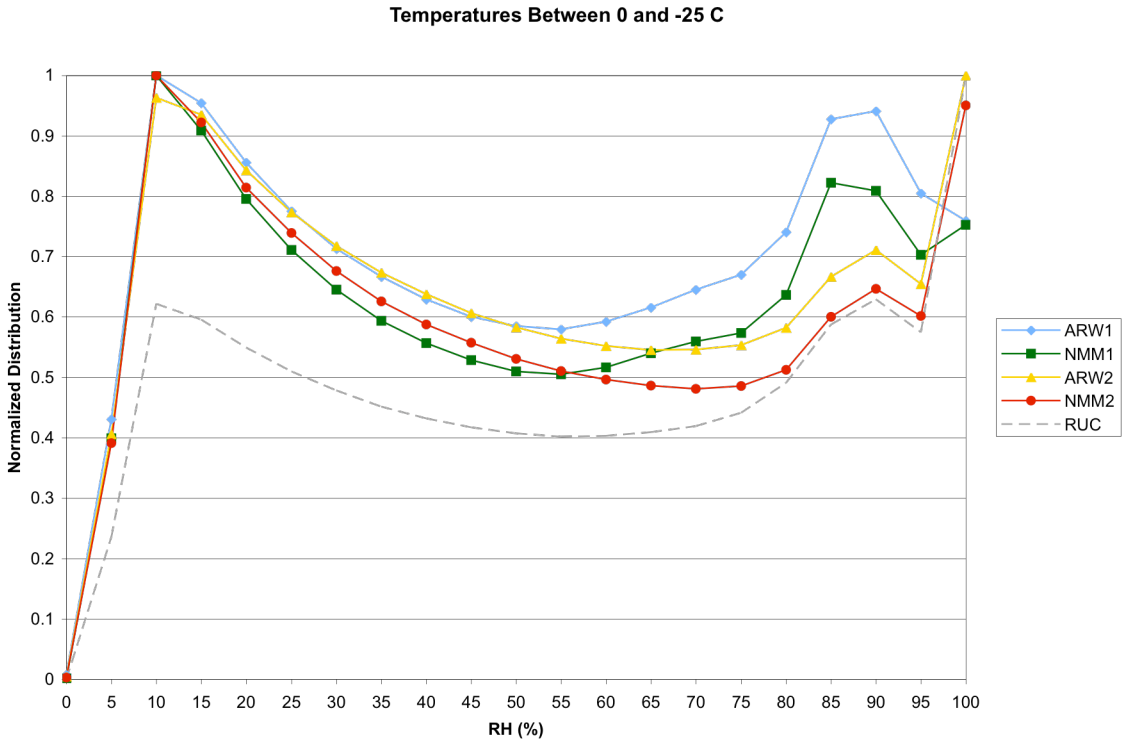
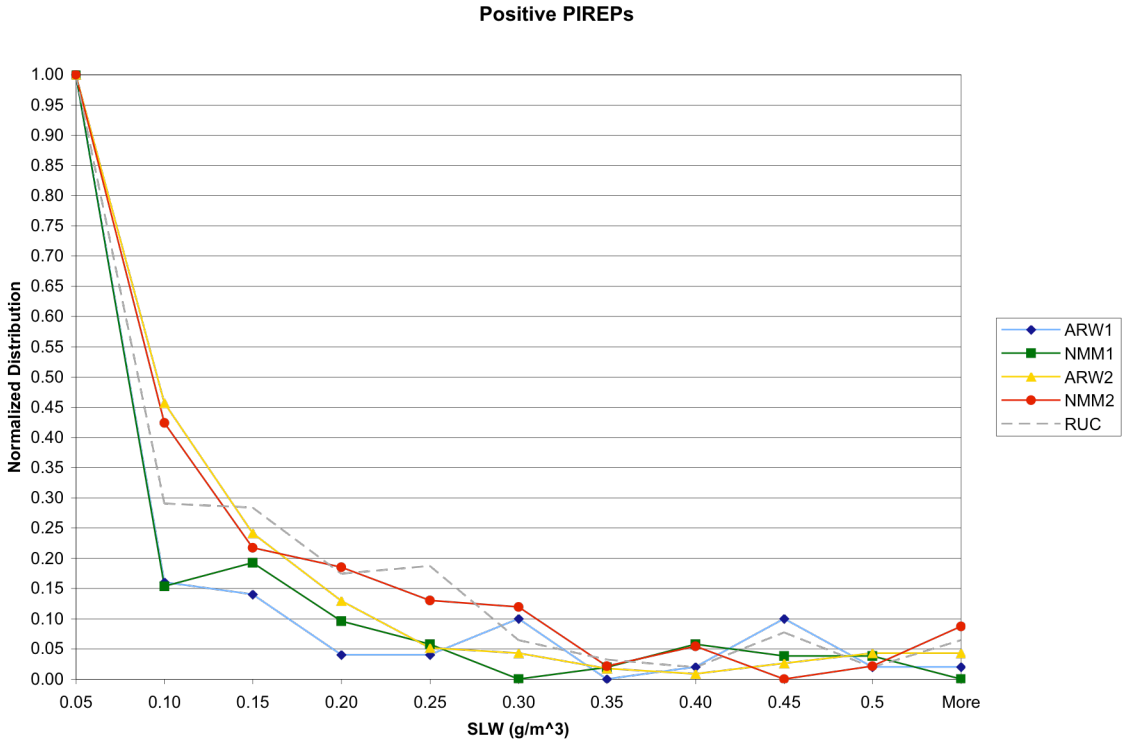


Fig. 2 – Same as Fig. 1, but for RH with respect to water.

SLW and TotC distributions are examined a bit differently because these fields are frequently equal to 0.0 in the overall grids. Recall that PODy for these grids was on the order of 0.23 for both cores run in phase 1 and 0.45 for both cores run in phase 2 (see Table 1). Thus, 77% and 55% of all positive icing PIREPs had SLW=0.0. While this is interesting in terms of how well the microphysics schemes can act as stand-alone predictors of in-flight icing, what is of equally great interest to the developers is the amount of SLW that is predicted, both at the location of positive PIREPs and overall. The positive PIREP plots are going to tend to be noisy because the total number of PIREPs within them is decreased dramatically. Figure 3 shows that relatively low SLW contents (SLWC <0.1 gm^{-3}) are predicted most frequently, regardless of core and phase. Core-to-core differences for phase 1 (Ferrier) appear to be negligible, but this is not the case for phase 2. Though the fields are noisy, the phase 2 distributions include ~500 positive PIREPs. They show that the NMM has a tendency to predict "moderate" SLWCs between 0.10 and 0.30 when positive PIREPs are present, while the ARW has a slight tendency toward SLWCs that are a bit more on the high (>0.4) or low (<0.2) end of the spectrum. These differences may not be significant because of the number of PIREPs. The overall distributions, which account for ~2 million grid points, show somewhat different trends, with the NMM producing SLWC in the highest bins nearly twice as frequently as the ARW, while the ARW has a fraction more SLWC in the much more commonly forecast lower bins.

TotC distributions show little or no core-to-core differences, regardless of phase, both versus positive PIREPs and overall (Fig. 4). Both cores tend to predict low amounts of total condensate, with the NMM producing such low amounts slightly more frequently than the ARW. Given the tendency for the phase 1 runs to predict ice condensate rather than liquid condensate at PIREPs, this may indicate that low amounts of ice condensate were commonly predicted at the locations of icing PIREPs.

Vertical velocity distributions do show some core-to-core differences, both versus positive PIREPs and overall (Fig. 5). While both models tend to be centric, with rather mild vertical velocities predicted most frequently, the ARW runs consistently had stronger upward and downward vertical velocities more frequently than the NMM (see cyan and yellow lines [ARW] compared to the green and red lines [NMM] in the +/- 0.3 to 1.0 range). A limited examination of a few case studies (see below) seems to indicate that this trend is seen because stronger vertical velocity couplets tend to be present in the ARW, while NMM vertical velocities appear to be more gentle and smooth. Interestingly, there appears to be very little difference in the phase-to-phase comparisons of this field, especially in the overall statistics. This clearly points to a core-to-core distinction. Note that the ARW distribution more closely matches that of the RUC, which tends to predict larger amplitude (positive and negative) vertical velocities even more frequently than the ARW.



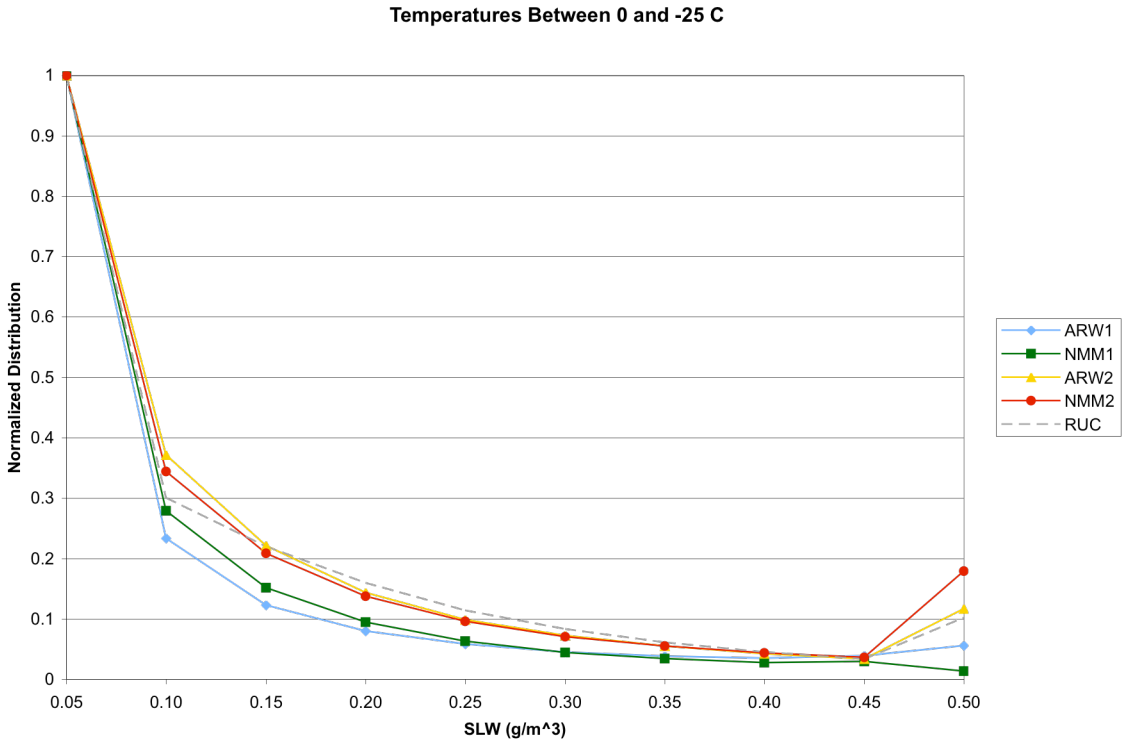


Figure 3 – Same as Fig. 1, but for SLWC.

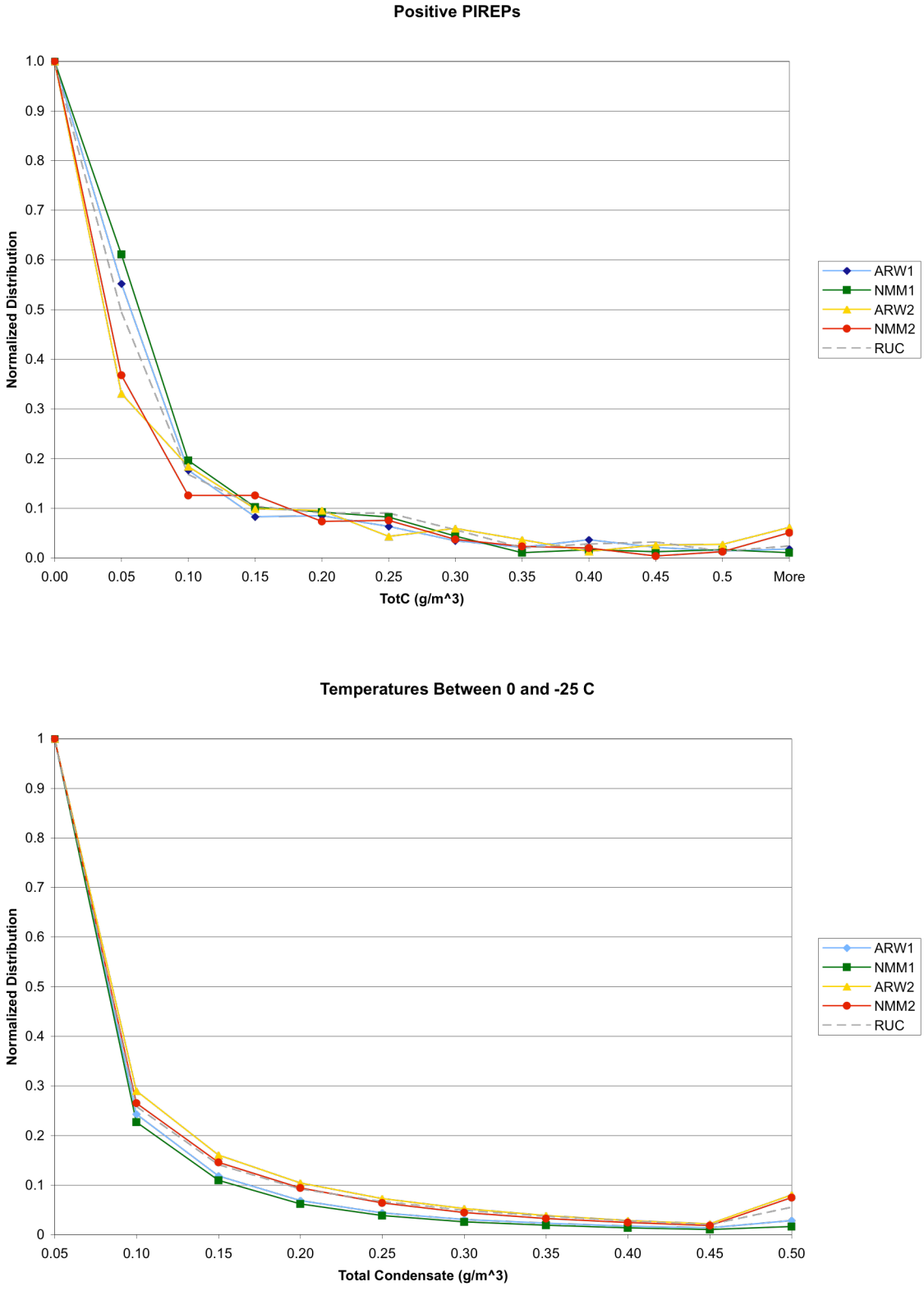


Figure 4 – Same as Fig. 1, but for total condensate (TotC).

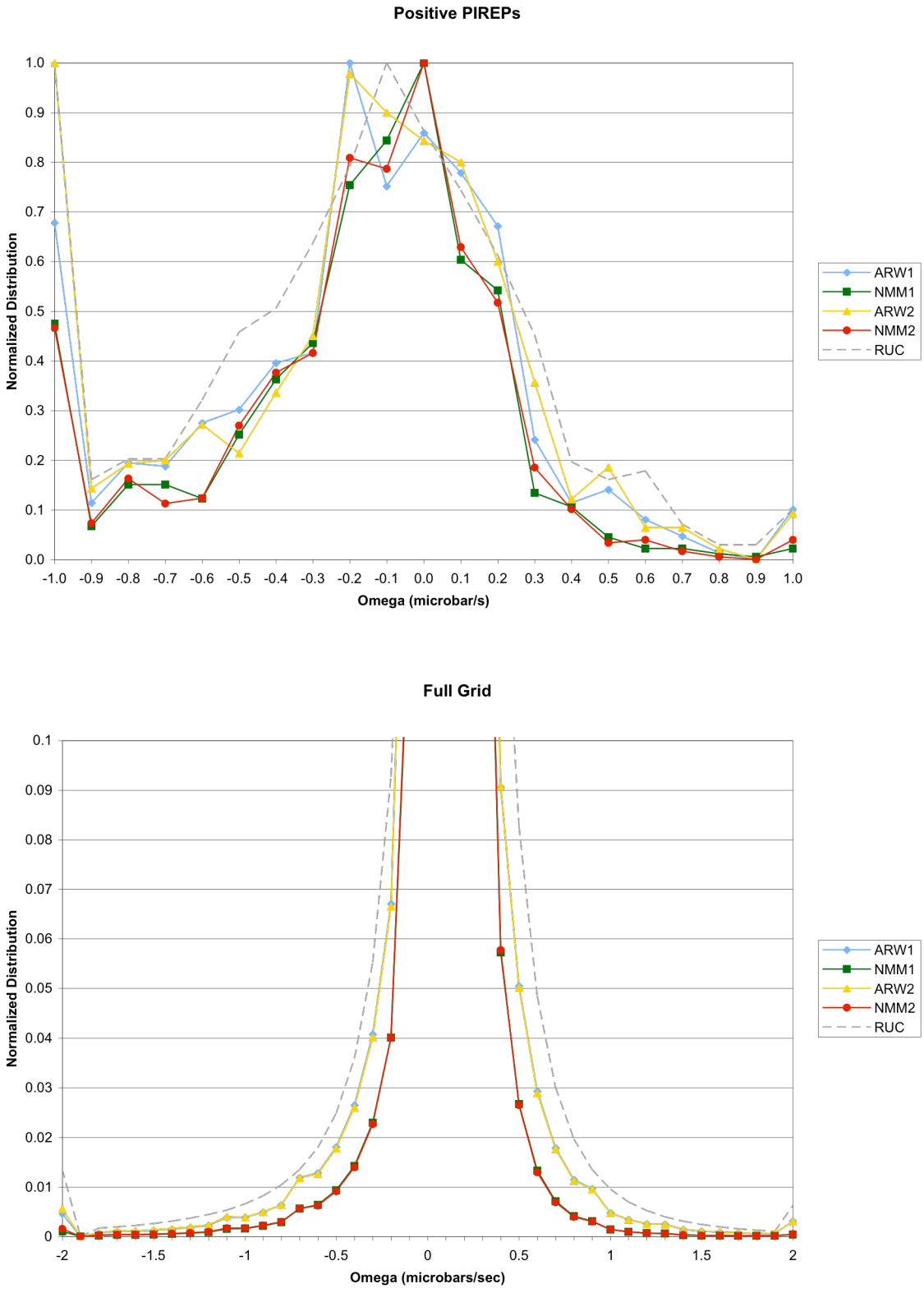


Figure 5 – Same as Fig. 1, but for omega.

An Example of the Differences in Scale Between the ARW and NMM

On 29 November 2005, a strong cold front passed through Ohio, reaching central Pennsylvania by 0000 UTC on 30 November. In its wake, cold advection, subfreezing temperatures and nearly saturated conditions were present across Ohio and the surrounding region (Fig. 6). This case was chosen because the level of interest is well above the effects of the topography, especially given the westerly component to the flow. Thus, the differences between the two runs should be dominated by differences in the models themselves.

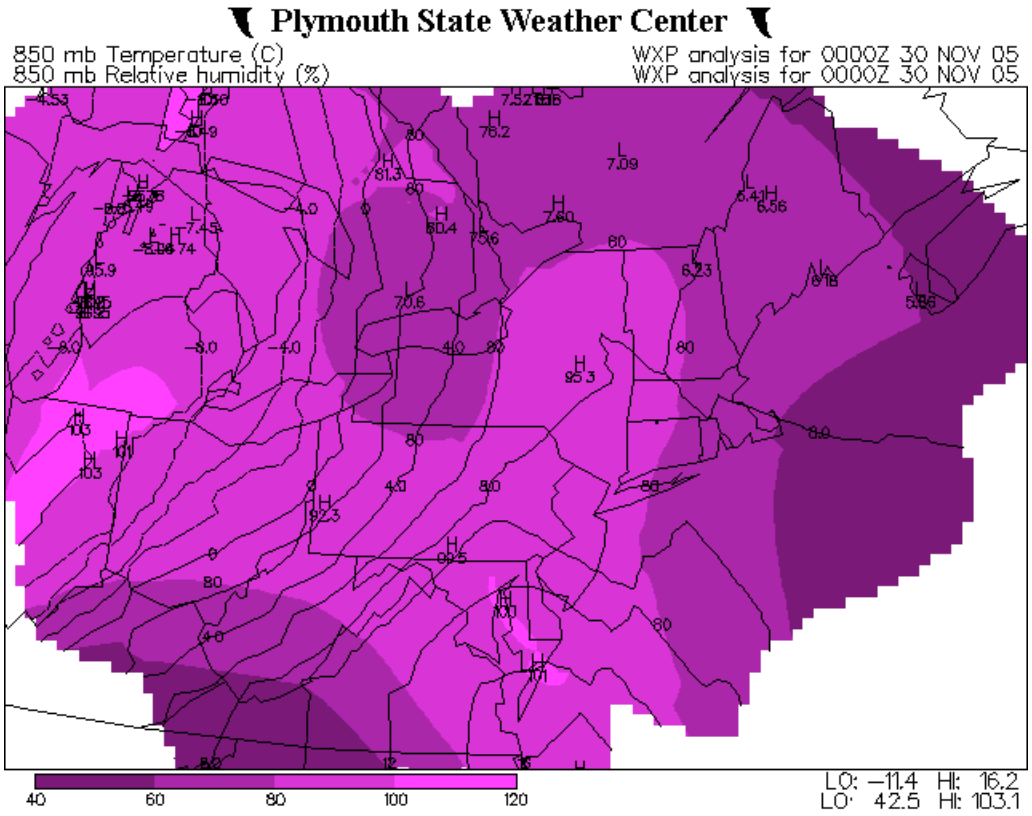


Figure 6 – Analysis of 850mb T (contours) and RH (filled) for 0000 UTC, 30 Nov 2005.

What makes this case of interest is the striking difference in the look, feel and scale of the vertical velocity, cloud liquid water content (supercooled across the northwestern ¾ of Ohio, as well as areas to the north and west – see the 0°C contour on Fig. 6) and the downstream effect on the FIP icing potential field. The comparison is made along constant k levels in the native model grids. Though the NMM and ARW grids do not match exactly, point comparisons of heights made across Ohio showed that they were typically within ~100m of one another, so the comparison is reasonable. Both the NMM and ARW plots are for 12h forecasts valid at 0000 UTC on 30 November. Phase 2 runs are used here because they are the only runs that produced SLW at this level for this case.

In what is generally a mild downward motion background field, the vertical velocities both show a small area of strong lift near the Ohio-Pennsylvania border (Fig. 7). A swath

of SLW is found in both models over this area, though their location, width and intensity differ somewhat. Differences between the models are much more evident over central Ohio and Lake Erie, where the NMM shows a patch of weak upward motion (-0.1 to -0.3 microbars/sec) over north Central Ohio that is surrounded by a broad area of neutral and weak to moderate downward motion. The ARW has a series of southwest-to-northeast striations of rather strong (near -1.0) surrounded by weak to moderate (+0.1 to +0.6) and a few pockets of relatively strong (+0.9) downward motion. The SLW fields reflect these differences, with a broad, north-south oriented swath of SLW in the NMM contrasted by striations of SLW that follow the rather sharp VV pattern in the ARW (Fig. 8).

The downstream effect on FIP is fairly evident in the icing potential fields created from the two model runs (Fig. 9). SLW and vertical velocity are used as "boosting factors" in FIP, taking initial estimates of icing potential and adjusting them downward (VV only) or upward (both fields) depending on where the magnitude and sign (VV only) of the values. SLW has the strongest effect that can only revise an initial icing potential estimate upward, while VV has a lesser effect that can revise it upward or downward. The southwest-northeast swaths of high icing potential (red) on the ARW FIP plot are a direct reflection of the VV and SLW pattern found in this model. Likewise, the adjacent moderate (relatively low) values of icing potential (yellow and orange). The NMM has a broader swath of high icing potential across central Ohio.

While the ARW run has a wave-like, more refined depiction of VV, SLW and icing potential, we have no way to verify that it existed on this day. Even if it existed, would the location and strength have been correct? Such differences in scale have shown up in several cases that we've examined, most times with the ARW showing finer-scale features than the NMM. This was especially true in areas with steep terrain, but since the underlying topography fields are somewhat different between the two models, we chose this more level terrain case to demonstrate the differences we see between the cores. The sharper gradients and larger magnitudes in the ARW VV field are likely what was reflected as a broader VV distribution for this core (Fig. 5). It is interesting to recall that the SLW distribution showed more occurrences of SLW in the highest bin for the NMM (Fig. 3). We speculate that such high values are more likely to occur when more gradual, long-lived and widespread upward motion is present, allowing the microphysics package adequate time to develop relatively high amounts of SLW.

We are encouraged to see the relatively fine scales that the ARW depicts, since we have found the icing conditions commonly vary on these and even finer scales in much of the research aircraft flight data we have analyzed. However, until we gain confidence that such scales are being represented accurately, we are not ready to take the fine scale information at face value and have it cascade readily into FIP and CIP. That said, in the meanwhile we may choose to damp the effects of such features while we investigate their validity. If they prove to be valid, we are excited at the possibility that finer-scale model features have the potential to more refined input as the primary ingredient to FIP and as background fields to CIP, perhaps coming a step closer to representing the scales upon which icing tends to occur and vary.

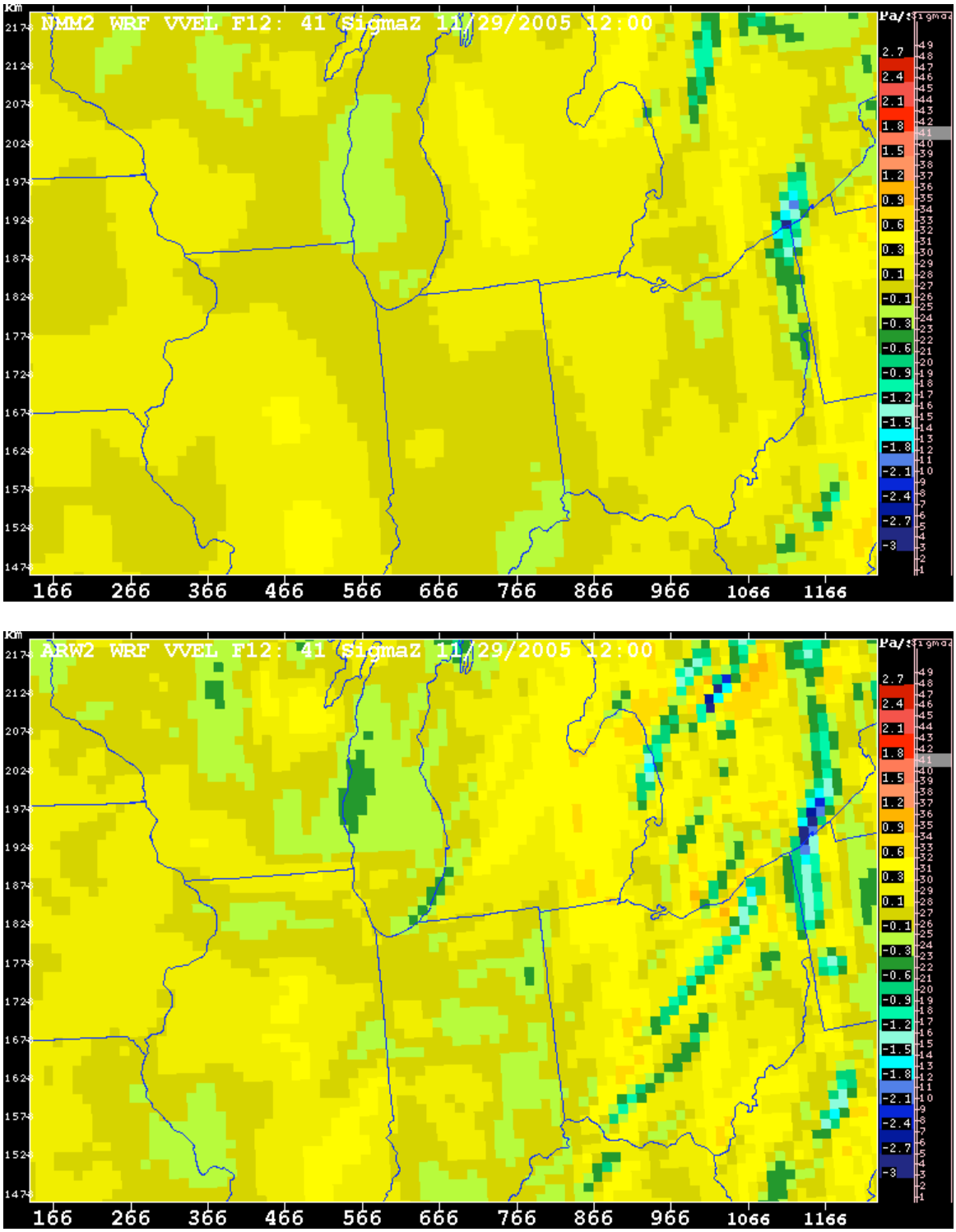


Figure 7 – NMM (top) and ARW (bottom) VV 12-h forecast fields for k=9 (shown as k=41 because of a reversal in the k-direction for the ingredient fields from the model), valid at 0000 UTC on 30 November 2005.

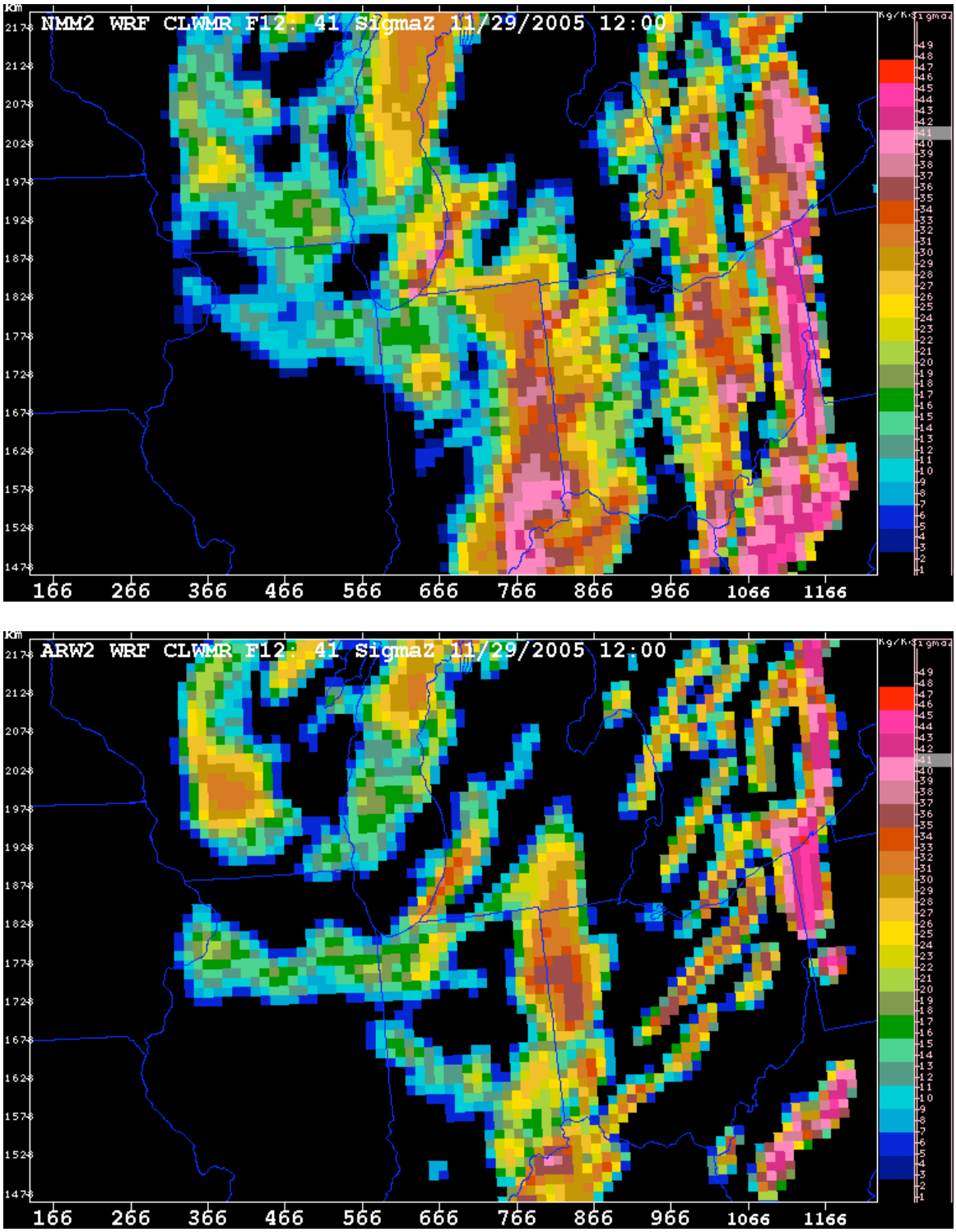


Figure 8 – Same as Fig. 7, but for cloud liquid water content.

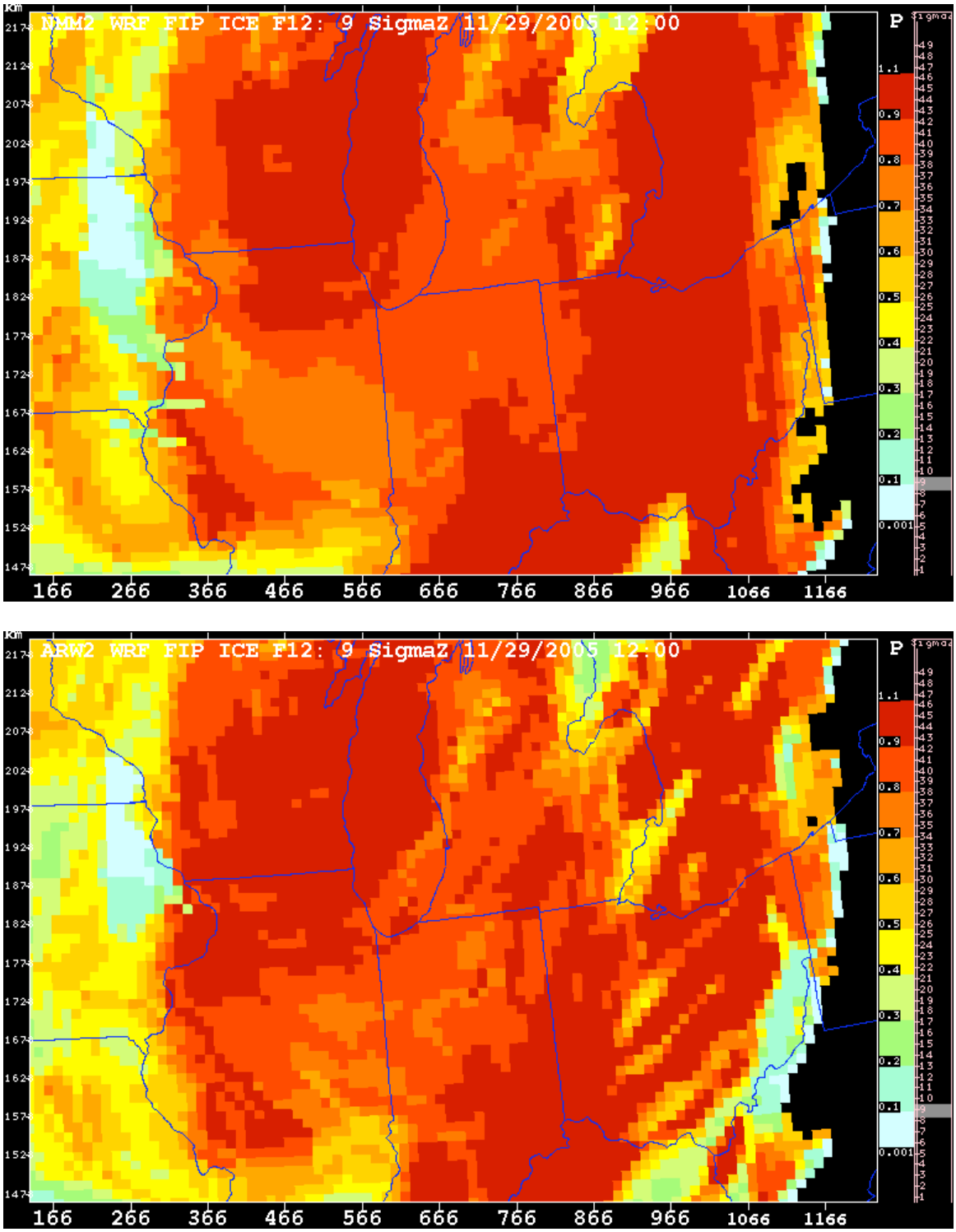


Figure 9 – Same as Fig. 7, but for FIP icing potential. The color scale goes from cool (low icing potential) to warm (high icing potential).